(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 869 478 A2

# (12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
**07.10.1998 Bulletin 1998/41**

(51) Int. Cl.[6]: **G10L 5/06**

(21) Application number: **98105750.8**

(22) Date of filing: **30.03.1998**

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE**
Designated Extension States:
**AL LT LV MK RO SI**

(30) Priority: **31.03.1997 JP 80547/97**

(71) Applicant: **NEC CORPORATION**
**Tokyo (JP)**

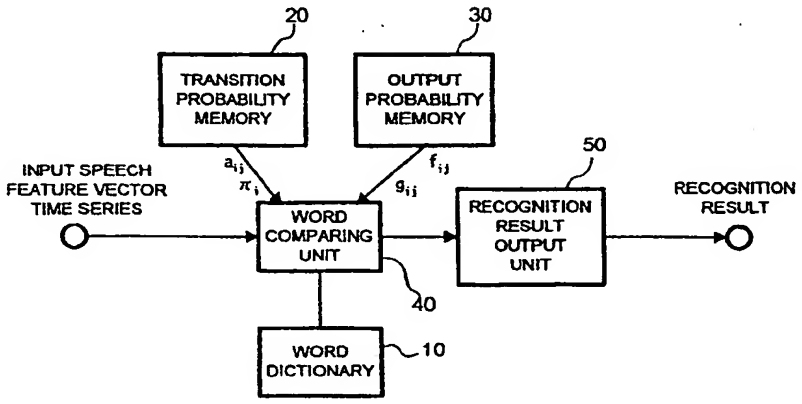
(72) Inventor: **Iso, Kenichi**
**Minato-ku, Tokyo (JP)**

(74) Representative:
**VOSSIUS & PARTNER**
**Siebertstrasse 4**
**81675 München (DE)**

## (54) Speech recognition method and apparatus

(57) Speaker independent speech recognition is made highly accurately without setting any recognition unit, such as triphone, and by taking environment dependency of phonemes into considerations. A word dictionary unit 10 stores phoneme symbol series of a plurality of recognition subject words. A transition probability memory unit 20 stores transition probabilities associated with N x N mutual state transitions of N states in a given order to one another. An output probability memory unit 30 stores phoneme symbol output probabilities and feature vector output probabilities associated with the respective state transitions. A work comparing unit 40 calculates probabilities of sets of unknown input speech feature vector time series and hypothetical recognition subject words. A recognition result output unit 50 provides a highest probability word among all the recognition subject words as a result of recognition.

**FIG. 1**




EP 0 869 478 A2

## Description

The present invention relates to speech recognition method and apparatus for recognizing unknown input speeches and, more particularly, to large vocabulary speech recognition method and apparatus which permit recognition of a large number of words.

For large vocabulary speech recognition, a method is extensively used, which have resort to triphone HMMs (Hidden Markov Models). Specifically, this method uses "triphone units" as recognition units, which are each prepared for adjacent phonemes present as a phoneme unit in a word (or sentence). The "triphone HMM" is detailed in "Fundamentals of Speech Recognition, Part I, Part II, NTT Advanced Technology Co., Ltd, ISBN-4-900886-01-7" or "Fundamentals of Speech Recognition, Prentice Hall, ISBN-0-13-055157-2".

In the speech recognition based on triphone HMMs, however, as many different HMMs as the cube of the number of different phonemes are involved, and it is difficult to accurately estimate all the triphone HMMs. To reduce the number of the different triphone HMMs, top-down or bottom-up clustering or the like is adopted, as detailed in the references noted above. Where the number of HMMs is reduced, however, it is no longer possible to guarantee the best fitness of the HMMs as such. In addition, such problem as having resort to intelligence concerening unreliable phonemes is posed.

An object of the present invention is to provide a method of and an apparatus for large vocabulary number speech recognition, which permits indefinite speaker's speech recognition highly accurately without setting triphones or like recognition units and by making even environment dependency of phonemes into considerations.

According to an aspect of the present invention, there is provided a speech recognition method of recognizing unknown input speech expressed as feature vector time series comprising the steps of storing phoneme symbol series of a plurality of recognition subject words, probabilities of N by N mutual state transitions of N states given sequential numbers to one another and phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions; calculating probabilities of sets of feature vector time series and unknown input speech and phone symbol series of provisional recognition subject words from an ergodic hidden Markov model; and outputting a maximum probability word among all the recognition subject words.

According to another aspect of the present invention, there is provided a speech recognition method of recognizing unknown input speech expressed as feature vector time series, comprising the sets of storing phone symbol series of a plurality of recognition subject words, probabilities of N by N mutual state transitions of N states given sequential numbers to one another, phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions and speaker's cluster numbers; and outputting a maximum probability word among all the recognition subject words.

According to other aspect of the speech recognition apparatus for recognizing unknown input speech expressed as feature vector time series comprising: a word dictionary unit for storing a plurality of phoneme symbol series of a plurality of recognition subject words; a transition probability memory unit for storing transition probabilities associated with N by N mutual state transitions of N states given sequential numbers to one another; an output probability memory unit for storing phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions; a word comparing unit for calculating probabilities of sets of feature vector time series of unknown input speech and phoneme symbol series of provisional recognition subject words; and a recognition result output unit for outputting maximum probability word among all the recognition subject words as recognition result.

According to still other aspect of the present invention, there is provided a speech recognition apparatus for recognizing unknown input speech expressed as feature vector time series comprising: a word dictionary unit for storing phone symbol series of a plurality of recognition subject words; a transition probability memory unit for storing transition probabilities associated with N by N mutual state transitions of N states given serial numbers to one another; an output probability memory unit for storing phone symbol output probabilities and feature vector output probabilities associated with the individual state transitions and speaker's cluster numbers; a word comparing unit for calculating probabilities of sets of feature vector time series of unknown input speech and phone symbol series of provisional recognition subject words; and a recognition result output unit for outputting a maximum probability word among all the recognition subject word and speaker's cluster numbers as recognition result.

The phoneme symbol is of a symbol by which a recognition subject word is defined absolutely or unanimously and is a syllable.

According to the present invention, speaker's cluster numbers associated with respective state transition may also be stored, and probabilities for time series of feature vector of unknown input speech, and sets of phoneme symbol series of provisionsl recognition subject words and provisional speaker's cluster number may be calculated, thereby outputting a maximum probability word among all the recognition subject words and speaker's cluster numbers.

The method of and apparatus for speech recognition according to the present invention is greatly different from the prior art method in that while in the prior art method feature vectors alone are provided in HMMs, according to the present invention phoneme symbols are also provided in HMM and speaker's cluster numbers are further provided in

the HMM. Furthermore, in the prior art a word HMM is constructed as reference pattern for each recognition subject word by connecting together triphone HMMs, whereas according to the present invention a single ergodic HMM is used as common reference pattern for all recognition subject words. That is, according to the present invention natural and common use of model parameter is realized.

Other objects and features will be clarified from the following description with reference to attached drawings.

Fig. 1 shows a block diagram of a speech recognition apparatus according to an embodiment of the present invention;

Fig. 2 shows probability of state transition from the state 1 to the state 2; and

Figs. 3 and 4 are flow chart illustrating a specific example of the routine.

Preferred embodiments of the present invention will now be described will now be described with reference to the drawings.

An embodiment of the speech recognition apparatus according to the invention is shown in Fig. 1. The speech recognition apparatus, which can recognize unknown input speech expressed as feature vector time series, comprises a word dictionary unit 10 for storing phoneme symbol series of a plurality of recognition subject words, a transition probability memory unit 20 for storing transition probabilities associated with N x N mutual state transitions of N states in a given order to one another, an output probability memory unit 30 for storing phoneme probabilities and feature vector output probabilities associated with the respective state transitions, a word comparing unit 40 for calculating probabilities of sets of unknown speech feature vector time series and hypothetical recognition subject words, and a recognition output unit 50 for providing a highest probability word among all the recognition subject words as a result of recognition.

The input speech is expressed as time series X

$$X = x_1 x_2 \cdots x_t \cdots x_T$$

of feature vectors $x_t$, where feature vector $x_t$ is, for instance, a 10-dimensional cepstrum vector, subscript t being number (natural number) representing sequential time.

In the word dictionary unit 10, phoneme symbol series of recognition subject words are stored. The phoneme symbol may sufficiently be of a Symbol unit less than a word, for instance a syllable, by which a recognition subject word can be defined absolutely or unanimously.

m-th recognition subject word is expressed as $w_m$, and its phoneme symbol series is expressed as

$$w_m = p_1 p_2 \cdots p_{Km}$$

where Km represents the length of the phoneme symbol series. The total number of phoneme symbols is $N_p$, and these phoneme symbols are given serial numbers.

TABLE 1

| Number | 1 | 2 | 3 | 4 | 5 | 6 | ... | Np |
|---|---|---|---|---|---|---|---|---|
| Phoneme Symbol | A | I | u | E | o | K | ... | ... |

For example, with a recognition subject word given by phonemes "akai", $p_1 = 1$, $p_2 = 6$, $p_3 = 1$, $p_4 = 2$, and Km = 4. The total number of recognition subject words is $N_w$. While in this embodiment phoneme symbols are used to express words, it is also possible to use other symbol systems such as syllables.

The HMM employed for speech recognition in this embodiment is ergodic HMM using ergodic Markov chain. The ergodic HMM is detailed in the literatures noted above. Fig. 2 is a view illustrating the ergodic HMM, which will now be described. Specifically, states 1 and 2 and all transitions associated with these states are shown. For example, $a_{12}$ in Fig. 2 represents the probability of state transition from the state 1 to the state 2. In the following, a case is considered, in which typically an ergodic HMM constituted by Ns states and mutual state transitions associated therewith is employed.

In the transition probability memory 20, probabilities of ergodic HMM state transitions are stored. The probabilities of transitions from i-th to j-th state are expressed as $a_{ij}$. The probabilities $a_{ij}$ meet conditions that their values are at least zero and that the sum of their values is 1, as shown by the following formula.

The initial probabilities of the states are

$$a_{ij} \geq 0,$$

$$\sum_{J=1}^{N_s} a_{ij} = 1$$

also stored in the transition probability memory 20.

The initial probabilities of i-th state are expressed as $\pi_i$. The initial probabilities $\pi_i$ meet the following conditions.

$$\pi_i \geq 0,$$

$$\sum_{i=1}^{N_s} \pi_i = 1$$

In the output probability memory 30, phoneme symbol output probabilities and feature vector output probabilities associated with state transitions are expressed as $f_{ij}(p)$ where p represents p-th phoneme symbols. Since the number of different phoneme symbols is $N_P$.

$$\sum_{p=1}^{N_p} f_{ij}(p) = 1$$

For example, $f_{ij}(1)$ represents the probability of output of phoneme symbol "a" in association with state transitions from i-th to j-th state.

Feature vector output probabilities associated with state transitions from i-th to j-th are expressed as $g_{ij}(x)$. The feature vector output probabiliies $g_{ij}(x)$ are herein expressed as multi-dimensional Gaussian distribution.

$$g_{ij}(x) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_{ij}|}} \exp[-(x - \mu_{ij})^t \sum_{ij}^{-1} (x - \mu_{ij})]$$

where D is the dimension number of the feature vectors, $\mu_{ij}$ is the mean vector, and $\Sigma_{ij}$ is the covariance matrix.

The word comparing unit 40 calculates probabilities (or likelihoods) of $N_w$ recognition subject words. The logarithmic value of probability $P(w_m, X)$ of m-th recognition subject word $w_m$ is calculated as follows. As noted before,

$$w_m = p_1 p_2 \cdots p_k \cdots p_{Km}, \text{ and}$$

$$X = x_1 x_2 \ldots x_t \ldots x_T.$$

The partial sum of logarithmic probabilities is defined as:

$$\phi_0(i,1) = \log[\pi_i],$$

$$\phi_0(i,k) = -\infty$$

$(1 < k \leq K_m)$

$$\phi_t(i,k) = \max\left[\max\left[\phi_{t-1}(j,k')\right] + \log\left[a_{ji}\right] + \log\left[f_{ji}(p_k)\right] + \log\left[g_{ji}(x_t)\right]\right]$$

$j = 1, \cdots, N_S$ $k' = k - 1, k$

$(1 \leq t \leq T, 1 \leq i \leq N_S, 1 \leq k \leq K_m)$

Using the above initialization and recurrence formula, the word comparing unit 40 calculates the partial sum $\phi_t(i, k)$ Of logarithmic probablities as three-dimensional array specified by three subscripts of t-th time, i-th state and k-th pho-

neme symbol for all times 1≦t≦T, all states 1≦i≦Ns and all phone symbols 1≦k≦Km in recognition subject word

From the partial sum $\phi_t(i, km)$ of logarithmic probabilities thus obtained, the logarithmic value of probablities $P(w_m, X)$ of m-th recognition subject word $w_m$ is obtained as:

$$\log[P(w_m,X)] = \max[\phi_T(i,K_m)]$$

$i = 1, \cdots, N_S$

The word comparing unit 40 calculates the logarithmic probabilities of all the recognition subject words. Figs. 3 and 4 are flow chart illustrating a specific example of the routine of the above process. In steps 101 to 108, the partial sum of logarithmic probabilities is initialized, in steps 109 to 133 the logarithmic value L of probability is calculated, and in step 134 the logarithmic value L is outputted. In the initialization routine, in step 102 i-th initial probability $\pi i$ is substituted into $\phi(0, i, 1)$ corresponding to t = 0, k = 1. For $\phi(0, 1, k)$ when k is at least 2, $-\infty$ is substituted in step 104. Since logarithmic probabilities are dealt with at this moment, $-\infty$ corresponds to anti-logarithm zero. Likewise, in sep 113 $-\infty$ is substituted into $\phi(t, i, k)$ as logarithm of anti-logarithm zero.

When the probabilities of all the recognition subject words have been obtained in the above way, the recognition result output unit 50 outputs word

$$W_{\hat{m}}$$

which gives the maximum probability among these probabilities as recognition result. That is,

$$\hat{m} = \arg\max_{m=1,\cdots,N_w}[\log[P(w_m, X)]]$$

While a preferred embodiment of the present invention has been described, it is by no means limitative. For example, while in the above embodiment the HMM output is provided by having feature vector output probabilities and phoneme symbol output probabilities associated with state transitions, it is possible to have also speaker's cluster number output probabilities associated with state transitions.

Where the speaker's cluster number output probabilities associated state transitions, the speaker's cluster number output probablities are expressed as $h_{ij}(q)$. Where the total number of speaker's clusters is $N_Q$, we have

$$\sum_{q=1}^{N_Q} h_{ij}(q) = 1$$

The speaker's cluster numbers are stored in the output probability memories 30. The initialization and recurrence formula noted above are expanded with the partial sum of logarithmic probabilities as a four-dimensional array as

$$\phi_0(i,1.q) = \log[\pi_i],$$

$$\phi_0(i,k,q) = -\infty,$$

$(1 < k \le K_m, 1 \le q \le Q)$

From the partial sum of logarithmic probabilities

$$\phi_t(i,k,q) = \max[\max[\varphi_{t-1}(j,k',q)]+\log[a_{ji}]+\log[f_{ji}(p_k)]+\log[g_{ji}(x_t)]+\log[h_{ji}(q)]]$$

$k' = k - 1, k$

$j = 1, \cdots, N_S$

$(1 \le t \le T, 1 \le i \le N_S, 1 \le k \le K_m, 1 \le q \le Q)$

From the partial sum of logarithm probabilities thus obtained, the logarithmic value of probability of recognition subject word wm is obtained as

$$\log[P(w_m,X)] = \max[\max \phi_T(i,K_m,q)]$$

$i = 1, \cdots, N_S$ $q = 1, \cdots, Q$

These calculations are executed in the word comparing unit 40.

The recognition result output unit 50 outputs a word of the maximum probability among all the recognition subject words and speaker's cluster numbers as recognition result.

By adding the speaker's cluster numbers to the ergodic HMM output, it is possibe to obtain speech recognition even with automatic determination of the optimum speaker character even in speaker independent speech recognition.

As has been described in the foregoing, according to the present invention by using a single ergodic HMM for outputting phoneme symbol series and feature vector series it is possible to realize a large vocabulary speech recognition apparatus, which does not require setting "triphones" or like recognition units and takes even environment dependency of phonemes into considerations. In addition, by adding speaker's cluster numbers to the output of the ergodic HMM output, it is possible to realize an apparatus, which can recognize speech with automatic determination of optimum speaker character even in speaker independent speech recognition.

changes in construction will occur to those skilled in the art and various apparently different modifications and embodiments may be made without departing from the scope of the present invention. The matter set forth in the foregoing description and accompanying drawings is offered by way of illustration only. It is therefore intended that the foregoing description be regarded as illustrative rather than limiting.

## Claims

1. A speech recognition method of recognizing unknown input speech expressed as feature vector time series comprising the steps of:

   storing phoneme symbol series of a plurality of recognition subject words, probabilities of N by N mutual state transitions of N states given sequential numbers to one another and phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions;
   calculating probabilities of sets of feature vector time series and unknown input speech and phone symbol series of provisional recognition subject words from an ergodic hidden Markov model; and
   outputting a maximum probability word among all the recognition subject words.

2. A speech recognition method of recognizing unknown input speech expressed as feature vector time series, comprising the steps of:

   storing phone symbol series of a plurality of recognition subject words, probabilities of N by N mutual state transitions of N states given sequential numbers to one another, phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions and speaker's cluster numbers; and
   outputting a maximum probability word among all the recognition subject words.

3. The speech recognition method as set forth in claim 1 or 2, wherein the phoneme symbol is of a symbol by which a recognition subject word is defined absolutely or unanimously.

4. The speech recognition method as set forth in claim 1 or 2, wherein the phoneme symbol is a syllable.

5. A speech recognition apparatus for recognizing unknown input speech expressed as feature vector time series comprising:

   a word dictionary unit for storing a plurality of phoneme symbol series of a plurality of recognition subject words;
   a transition probability memory unit for storing transition probabilities associated with N by N mutual state transitions of N states given sequential numbers to one another;
   an output probability memory unit for storing phoneme symbol output probabilities and feature vector output probabilities associated with the individual state transitions;
   a word comparing unit for calculating probabilities of sets of feature vector time series of unknown input speech and phoneme symbol series of provisional recognition subject words; and
   a recognition result output unit for outputting maximum probability word among all the recognition subject words as recognition result.

6. A speech recognition apparatus for recognizing unknown input speech expressed as feature vector time series comprising:

a word dictionary unit for storing phone symbol series of a plurality of recognition subject words;
a transition probability memory unit for storing transition probabilities associated with N by N mutual state transitions of N states given serial numbers to one another;
an output probability memory unit for storing phone symbol output probabilities and feature vector output probabilities associated with the individual state transitions and speaker's cluster numbers;
a word comparing unit for calculating probabilities of sets of feature vector time series of unknown input speech and phone symbol series of provisional recognition subject words; and
a recognition result output unit for outputting a maximum probability word among all the recognition subject word and speaker's cluster numbers as recognition result.

7. The speech recognition method as set forth in claim 5 or6, wherein the phoneme symbol is of a symbol by which a recognition subject word is defined absolutely or unanimously.

8. The speech recognition method as set forth in claim 5 or 6, wherein the phoneme symbol is a syllable.

FIG. 1

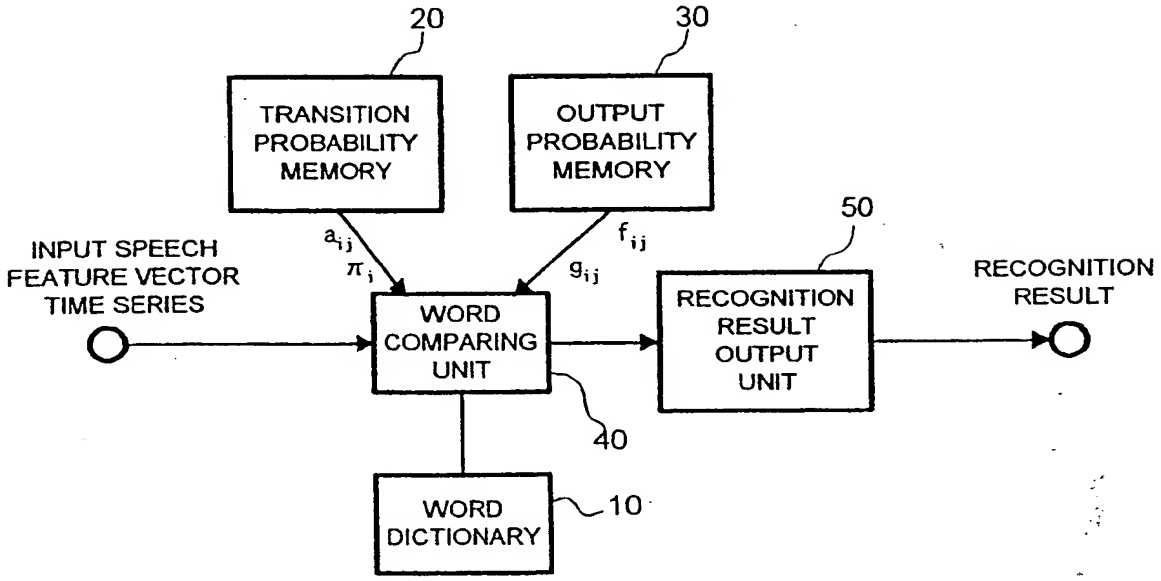
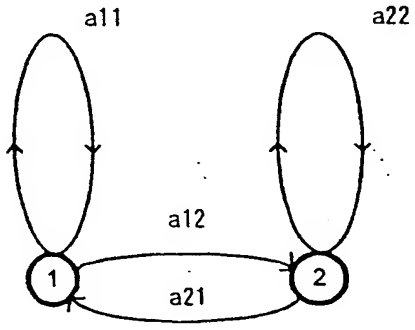
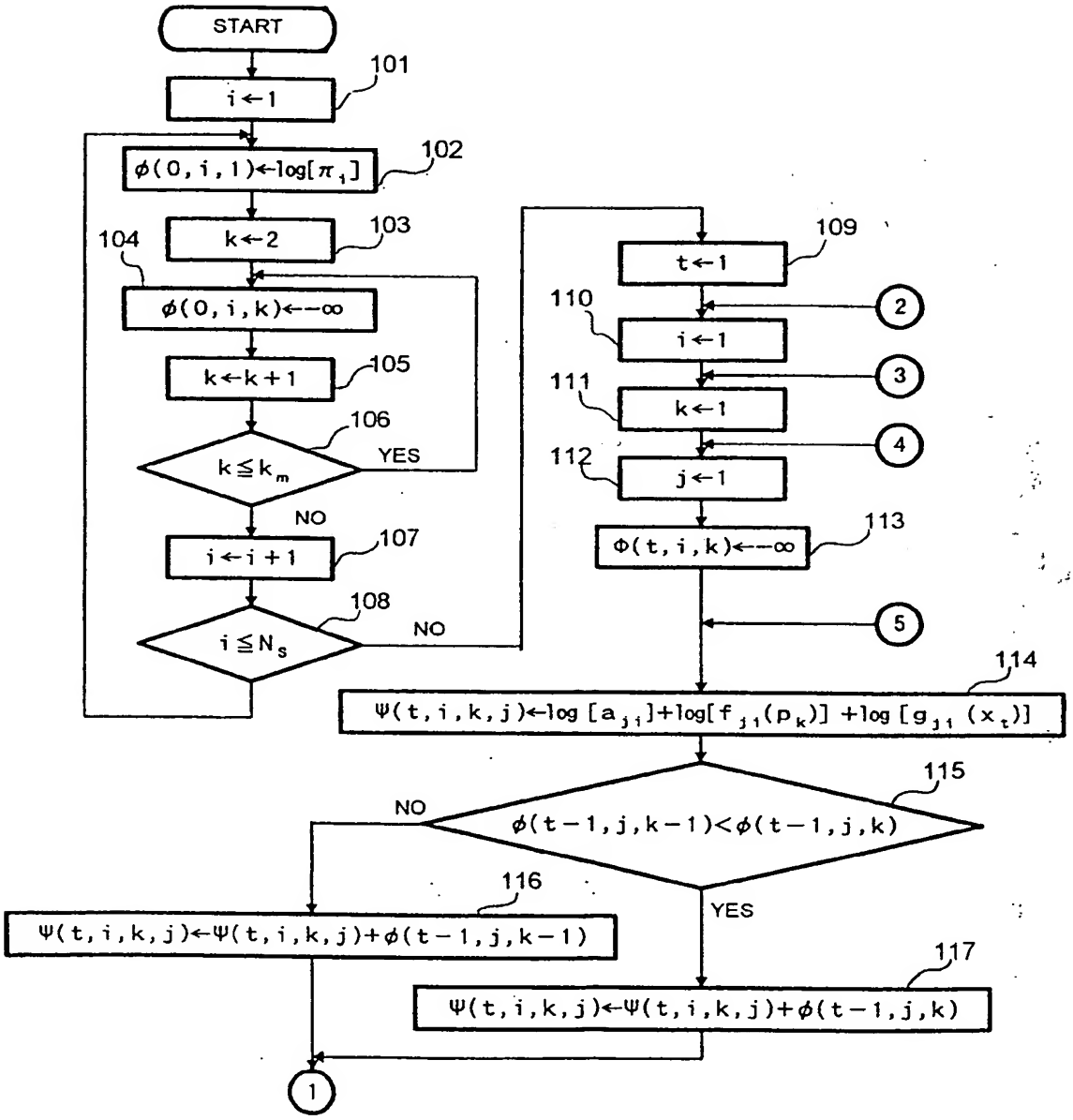

20
30
TRANSITION PROBABILITY MEMORY
OUTPUT PROBABILITY MEMORY
50
INPUT SPEECH FEATURE VECTOR TIME SERIES
$a_{ij}$
$\pi_i$
$f_{ij}$
$g_{ij}$
WORD COMPARING UNIT
RECOGNITION RESULT OUTPUT UNIT
RECOGNITION RESULT
40
WORD DICTIONARY
10

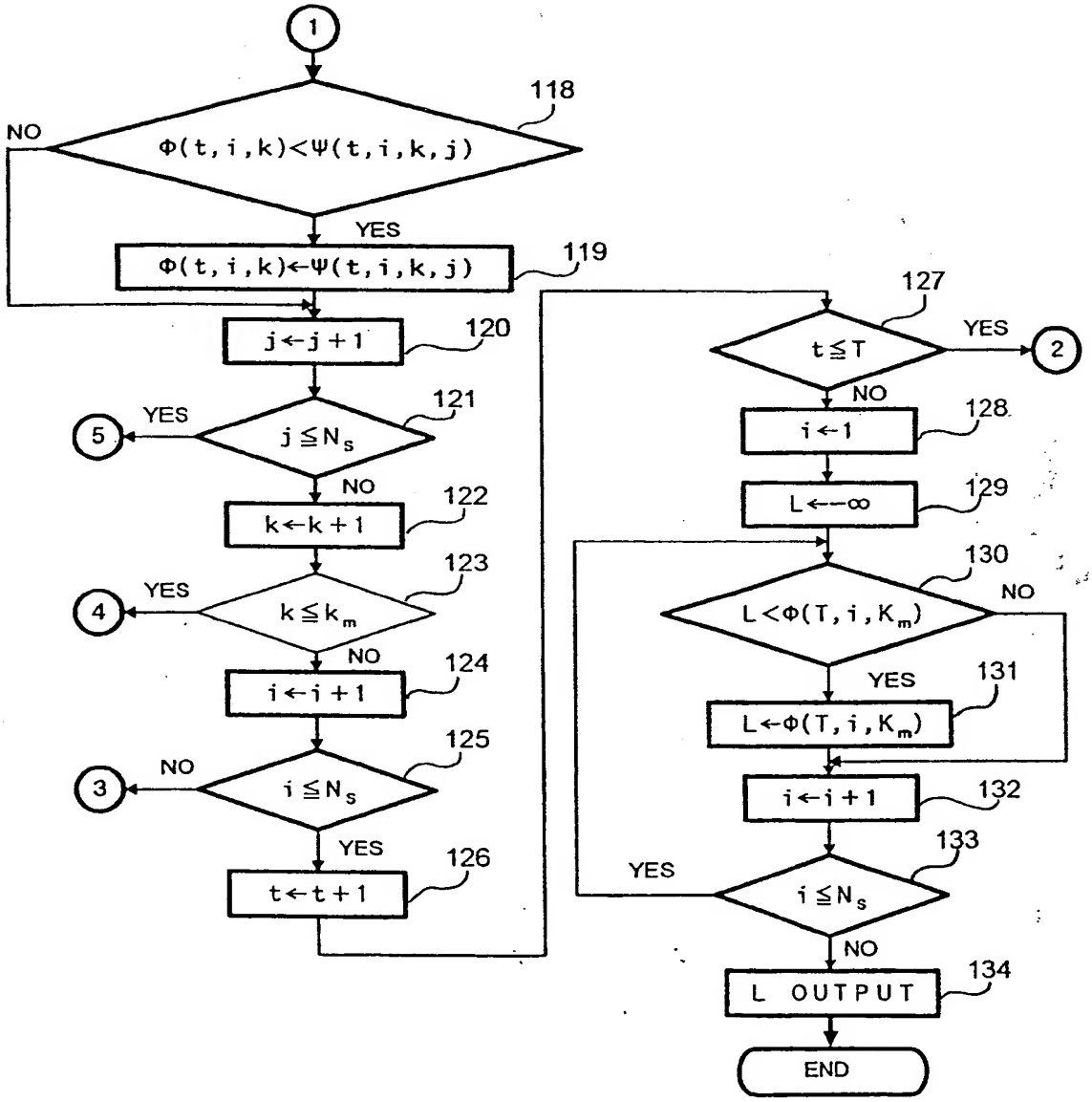
FIG. 2


a11
a22
a12
1
2
a21

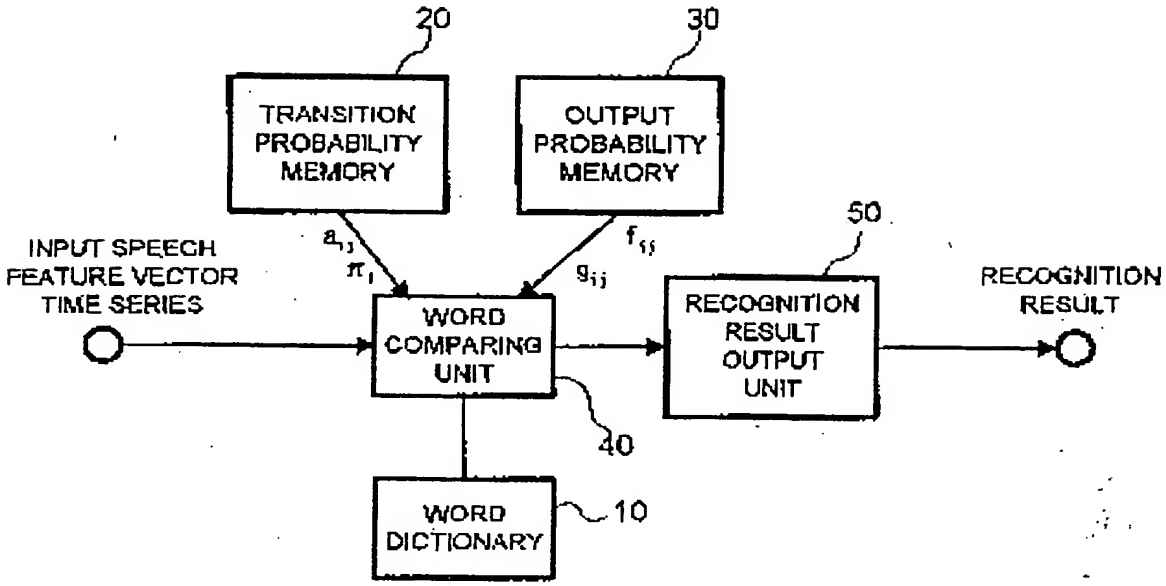
# FIG. 3


START
101
i←1
102
φ(0,i,1)←log[π_i]
103
k←2
104
φ(0,i,k)←−∞
105
k←k+1
106
k≦k_m
YES
NO
107
i←i+1
108
i≦N_s
NO
109
t←1
110
i←1
111
k←1
112
j←1
113
Φ(t,i,k)←−∞
2
3
4
5
114
Ψ(t,i,k,j)←log[a_ji]+log[f_ji(p_k)]+log[g_ji(x_t)]
115
φ(t−1,j,k−1)<φ(t−1,j,k)
NO
YES
116
Ψ(t,i,k,j)←Ψ(t,i,k,j)+φ(t−1,j,k−1)
117
Ψ(t,i,k,j)←Ψ(t,i,k,j)+φ(t−1,j,k)
1

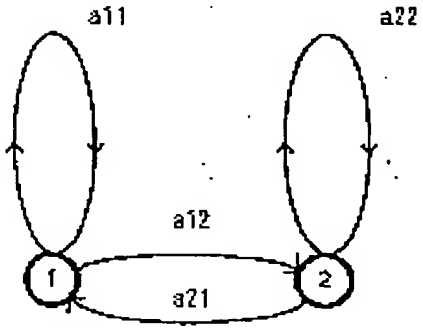
# FIG. 4



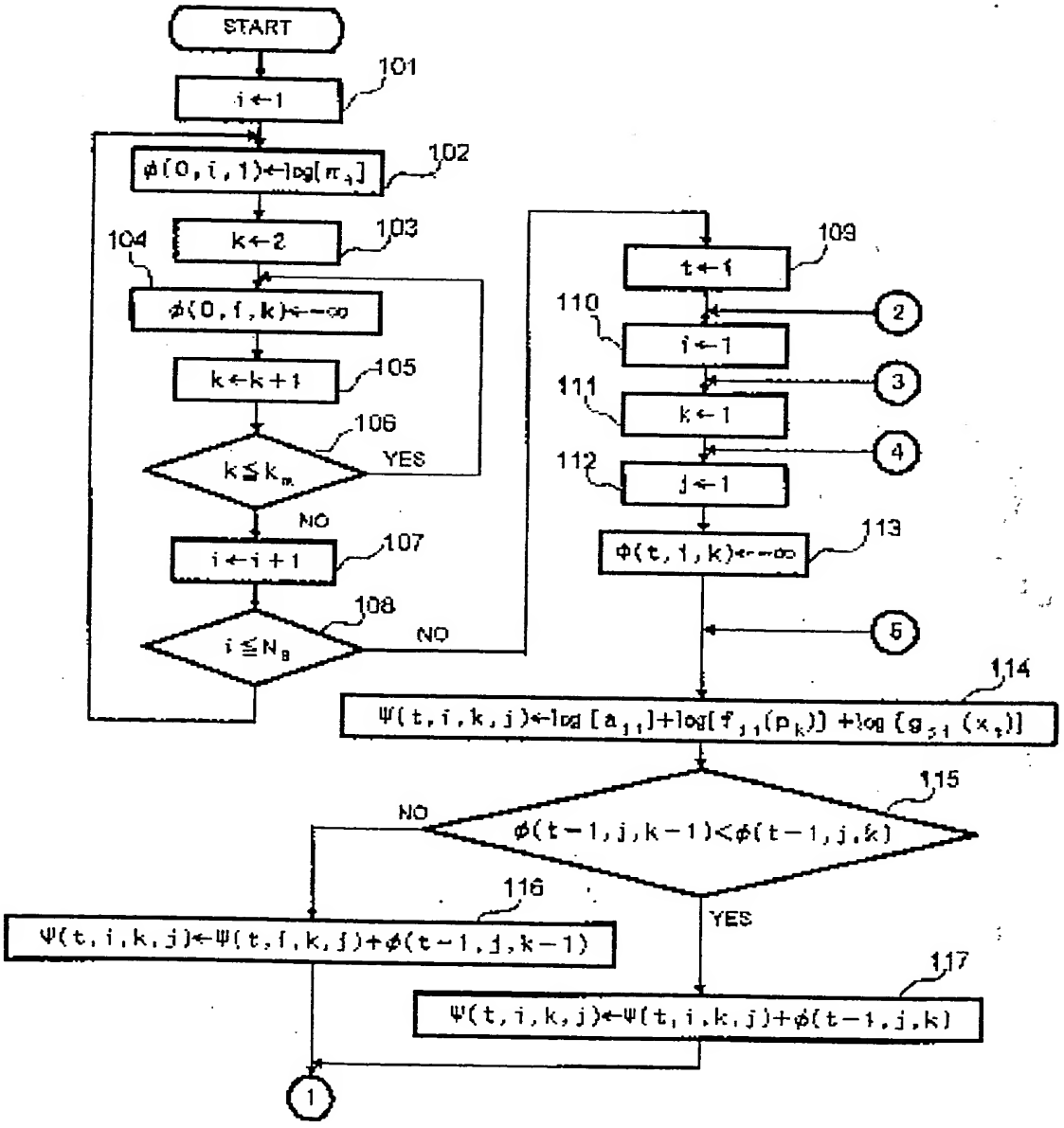
# FIG. 1


20
30
TRANSITION PROBABILITY MEMORY
OUTPUT PROBABILITY MEMORY
50
INPUT SPEECH FEATURE VECTOR TIME SERIES
$a_{ij}$
$\pi_i$
$f_{ij}$
$g_{ij}$
WORD COMPARING UNIT
RECOGNITION RESULT OUTPUT UNIT
RECOGNITION RESULT
40
WORD DICTIONARY
10

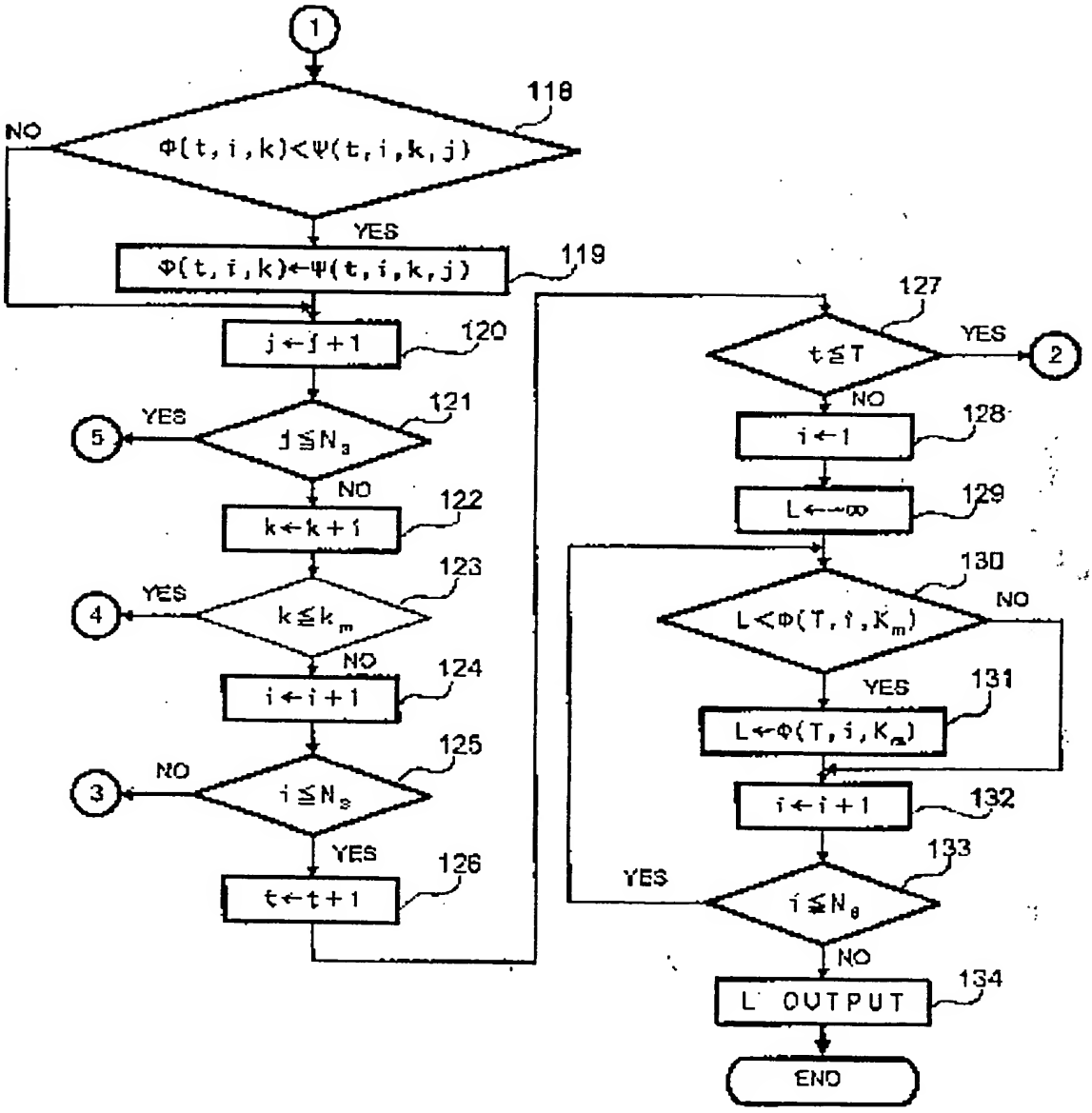
# FIG. 2


a11
a22
a12
a21
1
2

# FIG. 3


START
101
i←1
102
φ(0,i,1)←log[π_i]
103
k←2
104
φ(0,i,k)←−∞
105
k←k+1
106
k≦k_m
YES
NO
107
i←i+1
108
i≦N_s
NO
109
t←1
110
i←1
111
k←1
112
j←1
113
Φ(t,i,k)←−∞
2
3
4
5
114
Ψ(t,i,k,j)←log[a_j,i]+log[f_j,i(p_k)]+log[g_j,i(x_t)]
115
φ(t−1,j,k−1)<φ(t−1,j,k)
NO
YES
116
Ψ(t,i,k,j)←Ψ(t,i,k,j)+φ(t−1,j,k−1)
117
Ψ(t,i,k,j)←Ψ(t,i,k,j)+φ(t−1,j,k)
1

# FIG. 4


1
118
NO
Φ(t,i,k)<Ψ(t,i,k,j)
YES
Φ(t,i,k)←Ψ(t,i,k,j)
119
j←j+1
120
YES
5
121
j≦N_s
NO
k←k+1
122
YES
4
123
k≦k_m
NO
i←i+1
124
NO
3
125
i≦N_s
YES
t←t+1
126
127
t≦T
YES
2
NO
i←1
128
L←∞
129
130
L<Φ(T,i,K_m)
NO
YES
L←Φ(T,i,K_m)
131
i←i+1
132
YES
133
i≦N_g
NO
L OUTPUT
134
END

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 869 478 A3

# (12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
**26.05.1999 Bulletin 1999/21**

(43) Date of publication A2:
**07.10.1998 Bulletin 1998/41**

(21) Application number: **98105750.8**

(22) Date of filing: **30.03.1998**

(51) Int. Cl.[6]: **G10L 5/06**

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE**
Designated Extension States:
**AL LT LV MK RO SI**

(30) Priority: **31.03.1997 JP 80547/97**

(71) Applicant: **NEC CORPORATION**
**Tokyo (JP)**

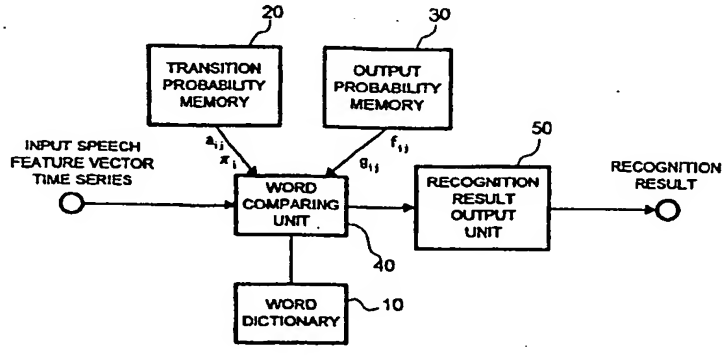
(72) Inventor: **Iso, Kenichi**
**Minato-ku, Tokyo (JP)**

(74) Representative:
**VOSSIUS & PARTNER**
**Siebertstrasse 4**
**81675 München (DE)**

## (54) Speech recognition method and apparatus

(57) Speaker independent speech recognition is made highly accurately without setting any recognition unit, such as triphone, and by taking environment dependency of phonemes into considerations. A word dictionary unit 10 stores phoneme symbol series of a plurality of recognition subject words. A transition probability memory unit 20 stores transition probabilities associated with N x N mutual state transitions of N states in a given order to one another. An output probability memory unit 30 stores phoneme symbol output probabilities and feature vector output probabilities associated with the respective state transitions. A work comparing unit 40 calculates probabilities of sets of unknown input speech feature vector time series and hypothetical recognition subject words. A recognition result output unit 50 provides a highest probability word among all the recognition subject words as a result of recognition.

FIG. 1




EP 0 869 478 A3

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 98 10 5750

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| X,P | EP 0 786 761 A (AT & T CORP) 30 July 1997<br>* abstract *<br>* page 4, line 26 - line 32 *<br>* page 5, line 2 - line 9 * | 1,5 | G10L5/06 |
| A | --- | 2,6 | |
| A | JP 06 266384 A (A T R JIDO HONYAKU DENWA KENKYUSHO:KK) 22 September 1994<br>* abstract * | 1,2,5,6 | |
| A | --- | | |
| A | L. RABINER, B. JUANG: "fundamentals of speech recognition"<br>1993 , PRENTICE HALL , ENGLEWOOD CLIFFS, NEW JERSEY 07632, USA XP002097960<br>* page 348, line 12 - line 22 * | 1,5 | |
| | ----- | | |
| | | | TECHNICAL FIELDS SEARCHED (Int.Cl.6) |
| | | | G10L |

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 25 March 1999 | Van Doremalen, J |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

## ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 10 5750

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

25-03-1999

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| EP 0786761 A | 30-07-1997 | US 5778341 A | 07-07-1998 |
| | | CA 2194617 A | 27-07-1997 |
| | | JP 9212188 A | 15-08-1997 |
| JP 06266384 A | 22-09-1994 | JP 1984185 C | 25-10-1995 |
| | | JP 7001435 B | 11-01-1995 |

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82